

B25

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
15 February 2001 (15.02.2001)

PCT

(10) International Publication Number  
**WO 01/11083 A2**

- (51) International Patent Classification<sup>1</sup>: C12Q 1/68 (74) Agent: CAMACHO, Jennifer, A.; Testa, Hurwitz & Thibault, LLP, High Street Tower, 125 High Street, Boston, MA 02110 (US).
- (21) International Application Number: PCT/US00/21763
- (22) International Filing Date: 9 August 2000 (09.08.2000) (81) Designated States (national): AU, CA, JP.
- (25) Filing Language: English
- (26) Publication Language: English (84) Designated States (regional): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
- (30) Priority Data: 09/371,991 11 August 1999 (11.08.1999) US Published:  
— Without international search report and to be republished upon receipt of that report.
- (71) Applicant: EXACT LABORATORIES, INC. (US/US); 63 Great Road, Maynard, MA 01754 (US).
- (72) Inventors: SHUBER, Anthony, P.; 11 Grant Street, Milford, MA 01757 (US). PIERCEALL, William; 17 Abbott Street, Wellesley, MA 02482 (US).
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*



WO 01/11083 A2

(54) Title: METHODS FOR DETECTING MUTATIONS USING PRIMER EXTENSION

(57) Abstract: Methods for detecting nucleotide deletions in biological samples are described. Methods of the invention are particularly useful for detecting deletions in regions of polynucleotide repeats. In particular, methods of the invention are useful to detect deletions at the BAT26 locus.

## METHODS FOR DETECTING MUTATIONS USING PRIMER EXTENSION

### FIELD OF THE INVENTION

The invention relates generally to methods for detecting nucleic acid mutations in biological samples, and more specifically to methods for detecting nucleic acid deletions or insertions using primer extension reactions.

### 5 BACKGROUND OF THE INVENTION

Numerous diseases are thought to be initiated by disruptions in genomic stability. For example, sickle cell anemia, phenylketonuria, hemophilia, cystic fibrosis, and various cancers have been associated with one or more genetic mutation(s). Increased knowledge of the molecular basis for disease has lead to a proliferation of screening  
10 assays capable of detecting disease-associated nucleic acid mutations.

One such method identifies a genomic region thought to be associated with a disease and compares the wild-type sequence in that region with the sequence in a patient sample. Differences in the sequences constitute a positive screen. See e.g., Engelke, *et al.*, *Proc. Natl. Acad. Sci.*, 85: 544-548 (1988). Such methods are time-  
15 consuming, costly, and often results in an inability to identify the mutation of interest. Thus, sequencing is not practical for large-scale screening assays.

A variety of detection methods have been developed which exploit sequence variations in DNA using enzymatic and chemical cleavage techniques. A commonly-used screen for DNA polymorphisms consists of digesting DNA with restriction  
20 endonucleases and analyzing the resulting fragments by means of Southern blots, as reported by Botstein *et al.*, *Am. J. Hum. Genet.*, 32: 314-331 (1980) and White *et al.*, *Sci. Am.*, 258: 40-48 (1988). Mutations that affect the recognition sequence of the endonuclease will preclude enzymatic cleavage at that site, thereby altering the cleavage pattern of the DNA. Sequences are compared by looking for differences in  
25 restriction fragment lengths. A problem with this method (known as restriction fragment length polymorphism mapping or RFLP mapping) is its inability to detect mutations that do not affect cleavage with a restriction endonuclease. One study reported that only

- 2 -

0.7% of the mutational variants estimated to be present in a 40,000 base pair region of human DNA were detected using RFLP analysis. Jeffreys, *Cell*, 18: 1-18 (1979).

Single-base mutations have been detected by differential hybridization techniques using allele-specific oligonucleotide probes. Saiki *et al.*, *Proc. Natl. Acad. Sci.*, 86: 6230-6234 (1989). Mutations are identified on the basis of the higher thermal stability of the perfectly-matched probes as compared to mismatched probes.

Disadvantages of this approach for mutation analysis include: (1) the requirement for optimization of hybridization for each probe, and (2) the nature of the mismatch and the local sequence impose limitations on the degree of discrimination of the probes. In practice, tests based only on parameters of nucleic acid hybridization function poorly when the sequence complexity of the test sample is high (*e.g.*, in a heterogeneous biological sample). This is partly due to the small thermodynamic differences in hybrid stability generated by single nucleotide changes. Therefore, nucleic acid hybridization is generally combined with some other selection or enrichment procedure for analytical and diagnostic purposes.

A number of detection methods have been developed which are based on template-dependent, primer extension. Those methods can be placed into one of two categories: (1) methods using primers which span the region to be interrogated for the mutation, and (2) methods using primers which hybridize upstream of the region to be interrogated for the mutation.

In the first category, U.S. Patent No. 5,578,458 reports a method in which single base mutations are detected by competitive oligonucleotide priming under hybridization conditions that favor the binding of a perfectly-matched primer as compared to one with a mismatch. U.S. Patent No. 4,851,331 reports a similar method in which the 3' terminal nucleotide of the primer corresponds to the variant nucleotide of interest. Since mismatching of the primer and the template at the 3' terminal nucleotide of the primer inhibits elongation, significant differences in the amount of incorporation of a tracer nucleotide result under normal primer extension conditions.

Methods in the second category are based on incorporation of detectable, chain-terminating nucleotides in the extending primer. Such single nucleotide primer-guided extension assays have been used to detect aspartylglucosaminuria, hemophilia B, and cystic fibrosis; and for quantifying point mutations associated with Leber Hereditary Optic Neuropathy. See, *e.g.*, Kuppuswamy *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 1143-

- 3 -

1147 (1991); Syvanen *et al.*, *Genomics*, 8: 684-692 (1990); Juvonen *et al.*, *Human Genetics*, 93: 16-20 (1994); Ikonen *et al.*, *PCR Meth. Applications*, 1: 234-240 (1992); Ikonen *et al.*, *Proc. Natl. Acad. Sci. USA*, 88: 11222-11226 (1991); Nikiforov *et al.*, *Nucleic Acids Research*, 22: 4167-4175 (1994). An alternative primer extension  
5 method involving the addition of several nucleotides prior to the chain terminating nucleotide has also been proposed in order to enhance resolution of the extended primers based on their molecular weights. See e.g., Fahy *et al.*, WO/96/30545 (1996).

Strategies based on primer extension require considerable optimization to ensure that only the perfectly annealed oligonucleotide functions as a primer for the  
10 extension reaction. The advantage conferred by the high fidelity of the polymerases can be compromised by the tolerance of nucleotide mismatches in the hybridization of the primer to the template. Any "false" priming will be difficult to distinguish from a true positive signal. The reaction conditions of a primer extension reaction can be optimized to reduce "false" priming due to a mismatched oligonucleotide. However, optimization  
15 is labor intensive and expensive, and often results in lower sensitivity due to a reduced yield of extended primer.

A number of mutations leading to various forms of cancer involve the deletion of multiple nucleotides from a genomic sequence. An example is the BAT26 segment of the MSH2 mismatch repair gene. The BAT26 segment contains a long poly-A tract. In  
20 certain cancers, a characteristic 5 base pair deletion occurs in the poly-A tract. Detection of that deletion may provide diagnostic information. Accordingly, the invention provides methods for detecting deletions in genomic regions, such as BAT26 and others, which may be associated with disease.

#### SUMMARY OF THE INVENTION

25 Methods of the invention provide assays for identification of a deletion in a genomic region suspected to be indicative of disease. In general, methods of the invention comprise annealing a primer upstream of a region in which a deletion is suspected to occur, extending the primer through the region, terminating extension at a known end-point, and comparing the length and/or weight of the extended primer with  
30 that of an extended primer from the corresponding wild-type (non-affected) region or a molecular weight standard (either known or run in parallel). In preferred embodiments, the extended primer is labeled downstream of the region suspected to be deleted. In a highly-preferred embodiment, the comparative length and/or molecular weight of the

- 4 -

extended primer is determined by gel electrophoresis or mass spectroscopy. Also in a highly-preferred embodiment, the region suspected to contain the deletion comprises a poly-nucleotide tract in which the deletion is suspected to occur, and the sequence immediately downstream of the region is known and does not repeat a nucleotide species present in the polynucleotide tract. Preferably, the polynucleotide tract comprise three, two, or preferably one, species of nucleotide as explained in detail below. Methods of the invention retain the specificity of primer extension assays while increasing their sensitivity by reducing background due to premature termination of the extension reaction. Therefore, methods of the invention provide a highly sensitive and highly specific assay for detecting a small amount of mutant nucleic acid in a heterogeneous sample of predominantly wild-type nucleic acid.

Methods of the invention provide screening assays for the detection of a deletion in a region of the genome comprising one, but no more than three, species of nucleotide, and that is characterized by having a sequence for primer hybridization immediately upstream, and a sequence immediately downstream that does not contain a nucleotide present in the region suspected to be deleted. In a preferred embodiment, methods of the invention comprise selecting a nucleic acid having a known wild-type sequence and having a region (the deletion of which is suspected in disease) comprising at most three different types of nucleotides; hybridizing an oligonucleotide primer, or pair of oligonucleotide primers, immediately upstream of the target region; extending the primer by using a polymerase in the presence of the nucleotide bases that are complementary to the nucleotide bases of the target region, thereby to form a primer extension product; further extending the primer extension product in the presence of a labeled nucleotide that is complementary to a nucleotide base downstream from the target region, but not complementary to a nucleotide base within the target region; and determining the size of the extension product compared to a standard (e.g., a wild-type product or a molecular weight standard).

In a preferred embodiment, the target region in which the deletion is suspected to occur is greater than five nucleotides long, and/or the deletion is great than three nucleotides long. In a preferred embodiment, the primer extension reactions are cycled by varying the reaction temperature through successive annealing, extending and denaturing temperatures. Preferably, the molecular weight standard is the wild-type extension product, or one that corresponds to the expected size for the extension product from the wild-type nucleic acid template. The presence of an extension product smaller than the molecular weight standard is indicative of the presence of a deletion in the target region of the nucleic acid template. In a preferred embodiment, the primer extension product is terminated by incorporating a terminator nucleotide that is

- 5 -

complementary to a nucleotide downstream from the target region in a wild type nucleic acid, but not complementary to any of the nucleotides of the target region. In a more preferred embodiment, the labeled nucleotide and the terminator nucleotide are the same. In an alternative embodiment, more than one labeled nucleotide base is  
5 incorporated into the extension product prior to incorporation of the terminator nucleotide. Preferably, the nucleotides incorporated during extension through the region suspected of containing a deletion are unlabeled. However, if those nucleotides are labeled, they are preferably distinguishable from the labeled nucleotide that is incorporated at the 3' end of the extension product.

10 In a preferred embodiment, methods of the invention comprise detecting a nucleic acid mutation in a biological sample, such as stool, urine, semen, blood, sputum, cerebrospinal fluid, pus, or aspirate, that contains a heterogeneous mixture of nucleic acid having a deletion in the target region and wild type nucleic acid. Such a deletion in the target region may be present in only about 1-5% of the nucleic acid  
15 molecules having the target region. To increase the sensitivity of the assay, the sample may comprise a polymerase chain reaction product. Method of the invention are particularly useful in analyzing a deletion in the target region that is indicative of the presence of cancerous or precancerous tissue in such a biological sample, including colorectal cancer or precancer detection in stool.

20 In another embodiment, methods of the invention comprise further extending the primer extension product in the presence of labeled and unlabeled nucleotides, the nucleotides being of the same type (i.e., A, T, C, or G) and being complementary to one or more nucleotide downstream from the target region but not complementary to a nucleotide within the target region. In one embodiment the ratio of the labeled  
25 nucleotide to unlabeled nucleotide is 1:1. Methods of the invention may also include incorporating more than one monomer of the labeled nucleotide or unlabeled nucleotide into the extension product.

In another embodiment, methods of the invention comprise detecting a deletion in a sample by selecting a nucleic acid with a known wild-type sequence and having a  
30 target region suspected of containing a deletion, wherein the target region contains at most three different types of nucleotide bases selected from the group consisting of dGTP, dATP, dTTP, and dCTP; hybridizing an oligonucleotide primer to a region upstream of said target region, in a nucleic acid sample; contacting said hybridized oligonucleotide primer with an extension reaction mixture comprising: i) nucleotides  
35 which are complementary to the nucleotides in the target region, ii) a labeled nucleotide which is complementary to a nucleotide found downstream from the target

- 6 -

region, but which is not complementary to any nucleotide base found within the target region, and iii) a terminator nucleotide which is complementary to a nucleotide found downstream from the target region, but which is not complementary to any nucleotide found in the target region; extending the hybridized oligonucleotide primer to generate a  
5 labeled extension product; and comparing the size of the labeled extension product from step d) to a molecular weight standard, wherein a labeled extension product smaller than the molecular weight standard is indicative of the presence of a deletion in the target region.

Methods of the invention are especially useful to detect indicia of cancer or  
10 precancer in a heterogeneous sample. Stool is a good example of a heterogeneous sample in which methods of the invention are useful. A typical stool sample contains patient nucleic acids, but also contains heterologous nucleic acids, proteins, and other cellular debris consistent with the lytic function of the various nucleases, proteinases and the like found in the colon. Under normal circumstances, stool solidifies as it  
15 proceeds from the proximal colon to the distal colon. As the solidifying stool passes through the colon, colonic epithelial cells are sloughed onto the stool. If a patient has a developing tumor or adenoma, cells from the tumor or adenoma will also be sloughed onto stool. Those cells, and/or their debris, will contain molecular indicia of disease (e.g., mutations or loss of heterozygosity). In the early stages of development, nucleic  
20 acid indicative of an adenoma or tumor comprise only about 1% of the nucleic acid in a voided stool. If left untreated, proportionately more disease-related nucleic acids are found in stool. Methods of the invention are useful for detecting early-stage lesions in heterogeneous samples such as stool. Methods of the invention result in a high degree of sensitivity and specificity for the detection of early-stage disease. Methods of the  
25 invention are especially useful in detecting, for example, adenomas in the colon. Adenomas are non-metastatic lesions that frequently have the potential for metastasis. If all adenomas in a patient are detected and removed, the probability of complete cure is virtually certain.

Deletions in the BAT26 locus of the MSH2 mismatch repair gene have been  
30 associated with colorectal cancer. Thus, in a highly-preferred embodiment, the region in which a deletion is suspected to occur is the BAT26 locus. That locus contains a polyA tract in which deletions have been associated with cancer or precancer. Use of methods of the invention on the BAT26 locus identifies the characteristic deletions by producing an extension product in affected DNA that is shorter than the expected wild-  
35 type extension product. Methods of the invention will be exemplified below using the BAT26 locus. However, methods of the invention are appreciated to be useful on any

- 7 -

genetic locus in which a deletion occurs. Especially useful loci are those indicative of disease, and especially cancer.

A detailed description of certain preferred embodiments of the invention is provided below. Other embodiments of the invention are apparent upon review of the  
5 detailed description that follows.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A shows BAT26 deletion detection using primer extension reactions that incorporate labeled bases before the 3' end of the extension product.

10 Figure 1B shows BAT26 deletion detection using primer extension reactions that incorporate labeled bases at the 3' end of the extension product.

Figure 2 shows deletion detection at the APC1309 locus.

#### DETAILED DESCRIPTION OF THE INVENTION

15 Methods of the invention provide highly sensitive assays for detecting the presence of mutations in nucleic acid samples. Methods of the invention are especially useful for detecting the presence of nucleic acid deletions and/or insertions in heterogeneous biological samples. In preferred embodiments, methods of the invention are useful to detect mutations at loci that are associated with a disease such as cancer.

20 In general, methods of the invention comprise identifying a target nucleic acid region that is suspected of being mutated, and interrogating the target region using a primer extension reaction. A primer is hybridized upstream of the target region and extended through the target region. The extension reaction is terminated at a site beyond the target region. The extension product is analyzed, and the size of the  
25 product is used as an indicator of the presence or absence of a mutation in the target nucleic acid region. In general, the presence of an extension product that is smaller than expected is indicative of the presence of a deletion in the target region. Conversely, the presence of a labeled extension product that is larger than expected is generally indicative of the presence of an insertion in the target region. However, the  
30 presence of a small or large labeled extension product can also be an indicator of a point mutation in the target region, as explained in greater detail in the following sections.



- 8 -

Methods of the invention are particularly useful when the target region contains a sequence that causes the extending polymerase to pause, stutter, or terminate prematurely. For example, regions containing nucleotide repeats such as a tract of a given nucleotide (such as the polyA tract at the BAT26 locus) dinucleotide or  
5 trinucleotide repeats. However, the invention is generally useful to detect mutations at loci having a known wild-type nucleic acid.

In a preferred embodiment, a primer is hybridized upstream of a target region that contains at most three different nucleotide bases. The hybridized primer is extended through the target region in the presence of unlabeled nucleotides that are  
10 complementary to nucleotides of the target region. The primer extension product is further extended in the presence of a labeled terminator nucleotide that is complementary to a nucleotide found downstream from the target region, but not found in the target region. An extension product is only labeled if the labeled terminator  
15 nucleotide is incorporated in the extension reaction. Consequently, an extension product is only labeled if it is extended through the target region, and along to the template nucleotide that is complementary to the labeled terminator nucleotide. Accordingly, prematurely terminated extension products are not labeled and do not interfere with the detection and analysis of labeled product by gel electrophoresis and autoradiography.

20 The present invention comprises embodiments wherein the primer is labeled, or wherein a labeled nucleotide is incorporated into the extension product before extension through the target region is complete, provided that an additional label is incorporated into fully extended products so that they can be distinguished from prematurely terminated extension products. In one embodiment, a primer is labeled  
25 with a first label, the labeled primer is hybridized upstream of the target region and extended through the target region, a second label is incorporated into the extension product downstream from the target region, and the extension reaction is terminated. Consequently, an extension product that terminates prematurely within the target region only contains the first label, whereas a fully extended product contains both the first and  
30 second label. Accordingly, diagnostically relevant extension products are those that contain both labels.

Methods of the invention also comprise assays in which the extension product is labeled and terminated in separate steps, after extension through the target region is complete. In one embodiment, a template nucleic acid comprises a target region  
35 consisting of a repeat of a first nucleotide base. Downstream from the target region is a second nucleotide base followed by a third nucleotide base. A primer is hybridized

- 9 -

upstream of the target region and extended through the target region in the presence of unlabeled nucleotides that are complementary to the first nucleotide. After extension through the target region is complete, the extension product is further extended in the presence of a labeled nucleotide that is complementary to the second nucleotide of the template. Finally, the labeled extension product is terminated via an extension reaction in the presence of a terminator nucleotide (such as a dideoxy nucleotide) that is complementary to the third nucleotide of the template. Other embodiments of this aspect of the invention are also described in the following sections.

Accordingly, an important aspect of the invention is a primer extension reaction wherein prematurely terminated extension products can be distinguished from complete extension products that have not undergone premature termination. Preferably, prematurely terminated extension products are not labeled, whereas complete extension products are detectably labeled. Figure 1 illustrates the usefulness of the invention in a deletion detection assay. The experimental details relating to Figure 1 are described in greater detail in Example 1. Figure 1 show that the invention provides an effective method for minimizing background when interrogating a target nucleic acid region suspected of containing a deletion. Figure 1A shows multiple samples that were analyzed by a primer extension assay that incorporated labeled nucleotides into the extension product upstream of the target region. In Figure 1B, the same samples were analyzed according to methods of the invention. Figure 1B does not contain the background of labeled prematurely terminated extension products that are seen in Figure 1A. Consequently, the presence of a deletion is clearly indicated in lane 7 of Figure 1B, whereas lane 7 of Figure 1A is more difficult to interpret.

Additional aspects of the invention are described in the following sections and illustrated by the Examples.

#### **Choosing the target region and the oligonucleotide primer**

Preferably, a locus associated with a disease such as cancer is chosen. Most preferably, a locus that is known to frequently exhibit one or more deletions is chosen. Useful loci include those containing at most 3 out of the 4 possible nucleotide bases. Preferably, a chosen locus comprises a polynucleotide region in which the deletion is suspected to occur. Once a locus is chosen, primers are designed or chosen to maximize specificity of binding to a nucleotide sequence immediately upstream of the region suspected of containing a deletion. The primer must hybridize immediately upstream of the region suspected of containing the deletion so that no labeled nucleotide is incorporated into the primer extension product.

- 10 -

### Sample preparation and hybridization

Methods of the invention are performed on any tissue or body fluid, including biopsy samples, and others having a high concentration of affected (i.e., mutated) cells or cellular debris. However, methods of the invention are particularly  
5 useful for detecting mutations in heterogeneous biological samples. A preferred sample is stool. For the analysis of stool samples, preferred methods of the invention comprise obtaining at least a cross-section or circumferential portion of a voided stool as taught in U.S. patent number 5,741,650, and co-pending, co-owned U.S. patent application serial number 09/059,718, both of which are incorporated by reference  
10 herein. While a cross-sectional or circumferential portion of stool is desirable, methods provided herein are conducted on random samples obtained from voided stool, which include smears or scrapings. Once obtained, the stool specimen is homogenized. A preferable buffer for homogenization is one that contains at least 16mM ethylenediaminetetraacetic acid (EDTA), as taught in co-pending, co-owned U.S. patent  
15 application serial number 60/122,177, incorporated by reference herein. It has been discovered that the use of at least 16mM EDTA, and preferably 100 mM EDTA greatly improves the yield of nucleic acid from stool. Thus, a preferred buffer for stool homogenization comprises phosphate buffered saline, 20-100 mM NaCl or KCl, at least 16mM EDTA, and optionally a detergent (such as SDS) and a proteinase (e.g.,  
20 proteinase K).

After homogenization, nucleic acid is preferably isolated from the stool sample. Isolation or extraction of nucleic acid is not required in all methods of the invention, as certain detection techniques can be adequately performed in homogenized stool without isolation of nucleic acids. In a preferred embodiment, however, homogenized  
25 stool is spun to create a supernatant containing nucleic acids, proteins, lipids, and other cellular debris. The supernatant is treated with a detergent and proteinase to degrade protein, and the nucleic acid is phenol-chloroform extracted. The extracted nucleic acids are then precipitated with alcohol. Other techniques can be used to isolate nucleic acid from the sample. Such techniques include hybrid capture, and  
30 amplification directly from the homogenized stool. Nucleic acids can be purified and/or isolated to the extent required by the screening assay to be employed.

Nucleic acids to be analyzed are chosen based upon known or suspected relationships between specific mutations and cancer or precancer. If desired, sequence-specific hybrid capture is used to isolate specific nucleic acids from the  
35 sample. Target nucleic acids may be analyzed by any method of the art. Examples of preferred methods include enumerative analysis of the loss of heterozygosity as taught

- 11 -

in U.S. patent number 5,670,325, incorporated by reference herein. Enumerative methods do not require knowledge of the sequence of a mutant nucleic acid. Rather such methods determine that there has been an alteration (deletion, substitution, addition, rearrangement, or other mutation) in a wild-type nucleic acid. The investigated  
5 loci are chosen based upon the likelihood of an alteration being associated with cancer or precancer. Enumerative methods compare the number in a sample of a wild-type nucleic acid known not to be altered in cancer or precancer with the number of a wild-type nucleic acid known or suspected to be altered in cancer or precancer. A statistically-significant difference in the two numbers indicates a positive screen.

10 **Primer extension, labeling and termination**

A hybridized primer is extended through the target region using known methods for primer extension, including extension using DNA polymerases. An extended primer preferably is labeled using a detectable label. Preferably, a labeled nucleotide is added to the extended primer once extension through the target region is complete. In a  
15 preferred embodiment, the labeled extension reaction is terminated at a predetermined position downstream from the target region. In a preferred embodiment, the labeling and termination steps are performed simultaneously. In one embodiment a labeled terminator nucleotide is incorporated into the extended primer downstream from the target region. Alternatively, the labeling and termination steps are performed  
20 separately. Preferably, the labeling and termination reactions are performed at about the same predetermined site downstream from the target region. If not, premature termination of a labeled extension product can interfere with the analysis of the results. Indeed, if a labeled primer extension product must be extended significantly in order to reach the predetermined termination site, then premature termination of the labeled  
25 extension product results in a shorter than expected labeled extension product. This short extension product may result in either a false positive indication of a deletion, or creates a background that interferes with the detection of a short extension product resulting from a deletion in the target region. Preferably the labeled base is also a terminator base. More preferably the labeled based is incorporated immediately  
30 upstream of the terminator base. Label is preferably a radioactive isotope. Alternatively a fluorescent tag, a molecular weight tag or other detectable label.

**Detection and analysis of the extension product**

While unlabeled primer extension products are contemplated, in preferred methods of the invention, only extension products that have been extended through the  
35 region suspected of containing a deletion are analyzed, because they are the only extension products that contain a detectable label. Extension products that terminate

- 12 -

prematurely within the region suspected of containing a mutation are not labeled and are not detected in the assay. Therefore, these premature extension products do not contribute to background noise that interferes with the analysis of the results.

Extended primer products are preferably detected using gel electrophoresis,  
5 mass spectroscopy, sequencing, and other methods for determining the differential length of two primers.

The following examples illustrate practice of the invention using deletion detection in the BAT26 and APC 1309 loci on samples prepared from stool specimens.

**EXAMPLE 1: Deletion detection at the BAT26 locus**

10 Experiments were conducted to demonstrate the usefulness of the invention to detect deletions in the BAT26 locus. The following experiment compares the specificity for detecting deletions at the BAT26 locus using primer extension reactions that incorporate label before extension through the target region versus primer extension reactions that incorporate label at the 3' end of the extension product.

15 The nucleic acid template was prepared as follows. Template nucleic acid containing the BAT26 locus was amplified by PCR. To each 50 ul PCR reaction tube, 40 ul of washed streptavidin coated Dynal beads were added and mixed by vortexing on a high setting for a few seconds. The mixture was incubated in a rack at room temperature for 15 minutes, and mixed by vortexing after 5 minutes and 10 minutes of  
20 the incubation period. The tube was placed in a magnetic tube holder, and the supernatant was removed. A 100 ul aliquot of 2X Binding & Wash buffer was added to each sample, and vortexed on a high setting for a few seconds. The tube was again placed in a magnetic tube holder and the supernatant was removed. A 100 ul aliquot of 0.1M NaOH was added to each tube, and mixed by vortexing on high for a few  
25 seconds. After a 5 minute incubation at room temperature, the tubes were placed in a magnetic tube holder, and the supernatant was removed. A further 100 ul of 0.1M NaOH was added, and vortexed for a few seconds. After placing the tube in a magnetic tube holder and removing the supernatant, 100 ul of 1X Binding & Washing buffer was added and vortexed for a few seconds on a high setting. The tube was  
30 placed in a magnetic tube holder, the supernatant was removed, and 100 ul of 1X TE pH 8.0 was added. The tube was vortexed on high for a few seconds, placed in a magnetic tube holder, and the supernatant was removed. The beads were resuspended in 100 ul of 0.1X TE pH 8.0 buffer by vortexing on high for a few seconds. The resulting samples were used in the assays, and may be stored at 4C for up to 1  
35 month.

- 13 -

In a first experiment, 5 ul of bead-bound PCR product was added to the following primer extension reaction mixture: 9.625 ul of sterile molecular biology grade diH<sub>2</sub>O, 2.5 ul of 10X Sequenase Buffer, 2.5 ul of 5 uM primer 1, 2.5 ul of 2 mM dATP, 2.5 ul of 50 uM ddGTP, 0.125 ul of 32P dTTP, and 0.25 ul of Sequenase.

- 5 The reaction mixture was cycled in an MJ Research Tetrad Thermalcycler according to the following temperature profile.

Temperature	Time	# Cycles
94 C	5 min	1
94 C	30 sec	
10 52 C	10 sec	30
72 C	10 sec	

- 4 C May be taken out of cycler immediately or after overnight run

- A 15 ul aliquot of formamide based stop solution was added to each sample and mixed by pipetting up and down 5 times. A 7 ul aliquot from each sample was analyzed using a 15% denaturing polyacrylamide gel with 7M Urea in 1X TBE running buffer. The gel was dried and analyzed using a Packard Instant Imager. Results are shown in Figure 1A. Lanes 1-8 are analyses of DNA obtained from patient stool samples. Lanes 9-14 are controls. Lane 9 contains no DNA template. Lanes 10, 13, and 14 contain, respectively, 0%, 1%, and 5% mutant DNA with a deletion within the polyA stretch of the BAT26 locus. Lanes 11 and 12 are no PCR controls.

- In a second experiment, 5ul of bead bound PCR product was added to the following primer extension reaction mixture: 7.125 ul of sterile molecular biology grade diH<sub>2</sub>O, 2.5 ul of 10X Sequenase Buffer, 2.5 ul of 5 uM primer 2, 2.5 ul of 2 mM dATP, 2.5 ul of 50 uM ddTTP, 2.5 ul of 0.1 uM dGTP, 0.125 ul of 32P dGTP, and 0.25 ul of Sequenase.

- The reaction mixture was exposed to the same temperature cycling as the reaction mixture in the first experiment, and the products were separated on a polyacrylamide gel under the same conditions. Lanes 1-14 of Figure 1B show results of this second experiment. The same nucleic acid templates were used in the reactions shown in lanes 1-14 of Figure 1A and lanes 1-14 of Figure 1B.

In the first experiment, shown in Figure 1A, the radioactive dGTP was incorporated into the primer extension product before it was extended through the

- 14 -

polyA stretch of the BAT26 locus. Primer 1 (5'- AGCCCTTAACCTTTTTCAGG-3', SEQ ID No: 1) used in the first experiment, hybridizes immediately upstream of a site where dTTP is incorporated (an A on the template strand). Accordingly, prematurely terminated extension products are labeled and appear as background in all of lanes 1-8.

In the second experiment, shown in Figure 1B, the radioactive dTTP was incorporated into the primer extension product after it was extended through the polyA stretch of the BAT26 locus. The 3' end of primer 2 (5'-GCCCTTAACCTTTTTCAGGT-3', SEQ ID NO: 2) used in the second experiment, includes the T that is immediately downstream from primer 1. Accordingly, in the second reaction, radioactive dTTP is only incorporated into the primer extension product after it has been extended through the polyA stretch. Furthermore, the extension reaction is also terminated close to the site of 32P dGTP incorporation. The second reaction mixture also contains ddTTP, and some of the extension products incorporate 32PdGTP followed by ddTTP at the T repeat downstream from the polyA stretch. Accordingly, in the second experiment, primer extension products that terminate prematurely within the polyA stretch are not labeled and are not seen as background in lanes 1-8, nor in control lanes 9-14. In Figure 1B, only lanes 6 and 7, and control lanes 13 and 14, contain short labeled primer extension product. The only samples that contained nucleic acid template having a deletion in the polyA stretch were the ones that were analysed in lanes 6, 7, 13, and 14. The sample of lane 6 was contaminated with a small amount of deleted template. The sample of lane 7 was from a patient with colon cancer associated with a deletion in the polyA stretch of the BAT26 locus. The samples of lanes 13 and 14 contained 1% and 5% mutant DNA, respectively.

A comparison of Figures 1A and 1B, shows that methods of the invention reduce the background of primer extension reactions. As a result, the analysis is much easier to interpret. Indeed, the presence of smaller than expected extension products in the second experiment is an indicator of the presence of mutant nucleic acid in the sample. In the first experiment, smaller than expected extension products are present in all reactions, and the analysis is more complicated.

In addition, methods of the invention, illustrated by the results of the second experiment, can be used to detect a very small amount of mutant nucleic acid in a heterogeneous sample containing mainly normal nucleic acid. The results shown in lanes 6 and 13 are the most striking. In Figure 1A, it is difficult to decide whether a deletion product is present in lanes 6 and 13. In contrast, a deletion product is clearly present in lanes 6 and 13 of Figure 1B.

- 15 -

Methods of the invention are particularly useful for analyzing loci such as BAT26, where a stretch of repeated nucleotide sequence interferes with efficient extension of DNA polymerase reactions. Premature termination of extension reactions is typically more frequent at such loci.

5   **EXAMPLE 2: Deletion detection at the APC 1309 locus**

A deletion of 5 nucleotides is often found at codon 1309 of the APC gene. The nucleotide sequence at this location is 5'-GAAAAGATT-3' (SEQ ID NO: 3) in the wild-type gene. Typical deletions consist of GAAAA (SEQ ID NO: 4), AAAAG (SEQ ID NO: 5), or AAAGA (SEQ ID NO: 6). To detect any of these deletions using a method of  
10   the invention, a 17 base oligonucleotide was designed to hybridize immediately upstream of the position of the first G (the G of the GAA codon above). Hybridized primer was extended in the presence of unlabeled dATP, unlabeled dGTP, and 33P-ddTTP. Accordingly, the extension product is only labeled if it is extended through the target region suspected of containing a deletion and the labeled ddTTP is incorporated.  
15   The expected wild-type product is 25 bases long, whereas any of the deletions described above generates a 20 base long extension product.

The extension reaction was performed on a duplicates of patient samples and the results are shown in Figure 2. Controls containing 0%, 1%, and 5% mutant nucleic acid were also analyzed that contained a 5bp deletion in BAT 26. The control results  
20   indicate that the presence of 1% mutant nucleic can be detected unambiguously. Both tests for patient #508 indicated the presence of a deletion at the 1309 locus. Patient 508 did indeed have colon cancer associated with a deletion at the 1309 locus.

In contrast, the results for patients without a deletion at the 1309 locus showed no background at the position characteristic of a deletion containing extension product.  
25   Accordingly, methods of the invention are useful for a simple test for the presence of a deletion at the 1309 locus.



- 16 -

## Claims

What is claimed is:

- 1 1. A method for detecting a nucleic acid insertion or deletion, the method comprising  
2 the steps of:
  - 3 a) selecting a nucleic acid having a known wild-type sequence and having a  
4 target region comprising at most three different types of nucleotide bases selected from  
5 the group consisting of dGTP, dATP, dTTP, and dCTP;
  - 6 b) contacting a sample with an oligonucleotide primer that is complementary to a  
7 portion of said nucleic acid immediately upstream of said target region;
  - 8 c) extending said primer in the presence of nucleotide bases that are  
9 complementary to the nucleotide bases of the target region, thereby to form a primer  
10 extension product;
  - 11 d) extending the primer extension product in the presence of a labeled  
12 nucleotide complementary to a nucleotide base downstream from the target region in a  
13 said nucleic acid, wherein said labeled nucleotide is not complementary to any of the  
14 nucleotide bases of the target region; and,
  - 15 e) comparing the size of the labeled extension product obtained in step d) to a  
16 standard, wherein a labeled extension product smaller than the standard is indicative of  
17 the presence of a deletion in the target region and a labeled extension product larger  
18 than the standard is indicative of the presence of an insertion in the target region.
- 1 2. The method of claim 1, further comprising the step of terminating the primer  
2 extension product by incorporating a terminator nucleotide in said product that is  
3 complementary to a nucleotide downstream from the target region in a wild type nucleic  
4 acid, wherein said terminator nucleotide is not complementary to any of the nucleotides  
5 of the target region.
- 1 3. The method of claim 2, wherein the labeled nucleotide and the terminator nucleotide  
2 are the same.
- 1 4. The method of claim 2, wherein more than one labeled nucleotide is incorporated  
2 into the extension product prior to incorporation of the terminator nucleotide.
- 1 5. The method of claim 1 or 2, wherein the nucleotides of step c) are unlabeled.
- 1 6. The method of claim 1 or 2, wherein the labeling reaction of step d) is performed in  
2 the presence of labeled nucleotide and unlabeled nucleotide of the same type.

- 17 -

- 1 7. The method of claim 6, wherein the ratio of labeled nucleotide base to unlabeled  
2 nucleotide base is 1:1.6 (unlabeled:labeled).
- 1 8. The method of claim 6, wherein more than one nucleotide from step d) is  
2 incorporated into an extension product.
- 1 9. The method of claim 8, wherein only one of the incorporated nucleotides is labeled.
- 1 10. The method of claim 1 or 2, wherein said biological sample contains a  
2 heterogeneous mixture of mutant nucleic acid having a deletion in the target region and  
3 wild type nucleic acid with no deletion in the target region.
- 1 11. The method of claim 10, wherein a deletion in the target region is present in from  
2 about 1% to about 5% of the nucleic acid molecules containing the target region.
- 1 12. The method of claim 1 or 2, wherein said sample is a biological sample.
- 1 13. The method of claim 12, wherein said biological sample is selected from the group  
2 consisting of stool and homogenized stool.
- 1 14. The method of claim 12, wherein said biological sample is selected from the group  
2 consisting of urine, semen, blood, sputum, cerebrospinal fluid, pus, and aspirate.
- 1 15. The method of claim 12, wherein a deletion in the target region is indicative of the  
2 presence of cancerous or precancerous tissue in the biological sample.
- 1 16. The method of claim 1 or 2, wherein said target region is the polyA tract at the  
2 BAT26 locus.
- 1 17. The method of claim 1 or 2, wherein the presence of a deletion in said target region  
2 is associated with the presence of a mutation at a separate genetic locus.
- 1 18. The method of claim 17, wherein said separate genetic locus is a genetic locus  
2 associated with cancer or precancer.
- 1 19. The method of claim 18, wherein said genetic locus is selected from the group  
2 consisting of APC, DCC, P53, and RAS.
- 1 20. The method of claim 15, wherein said cancerous or precancerous tissue is of  
2 colorectal origin.
- 1 21. The method of claim 18 wherein said cancer or precancer is a colorectal cancer or  
2 precancer.
- 1 22. The method of claim 1 or 2, wherein a pair of oligonucleotide primers is hybridized  
2 immediately upstream of the target region.

- 18 -

1 23. The method of claim 1 or 2, wherein said extension reactions are catalyzed by a  
2 thermostable polymerase.

1 24. The method of claim 23, wherein more than one cycle of extension reactions is  
2 performed by cycling the reaction temperature through successive annealing,  
3 extending, and denaturing temperatures.

1 25. The method of claim 1 or 2, wherein the sample contains a polymerase chain  
2 reaction product.

1 26. A method for detecting a nucleic acid deletion in a sample, the method comprising  
2 the steps of:

3 a) selecting a nucleic acid with a known wild-type sequence and having a target  
4 region suspected of containing a deletion, wherein said target region contains at most  
5 three different types of nucleotides selected from the group consisting of dGTP, dATP,  
6 dTTP, and dCTP;

7 b) hybridizing an oligonucleotide primer to a region upstream of said target  
8 region, in a nucleic acid sample;

9 c) contacting said hybridized oligonucleotide primer with an extension reaction  
10 mixture comprising:

11 i) the nucleotides that are complementary to the nucleotides in the target  
12 region,

13 ii) a labeled nucleotide that is complementary to a nucleotide found  
14 downstream from the target region, but is not complementary to any nucleotide found  
15 within the target region, and

16 iii) a terminator nucleotide that is complementary to a nucleotide found  
17 downstream from the target region, but is not complementary to any nucleotide found in  
18 the target region;

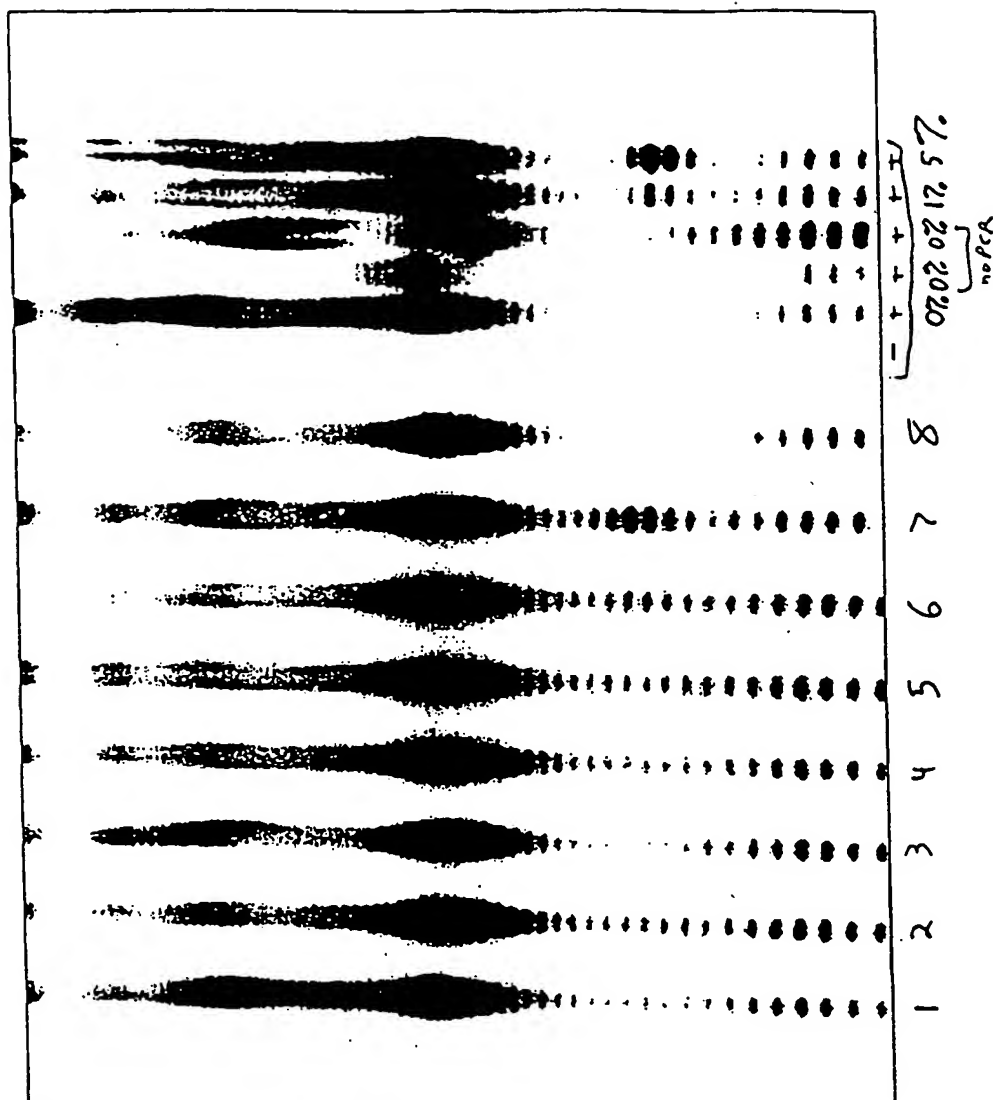
19 d) extending the hybridized oligonucleotide primer to generate a labeled  
20 extension product; and

21 e) comparing the size of the labeled extension product from step d) to a  
22 standard, wherein a labeled extension product smaller than the is indicative of the  
23 presence of a deletion in the target region, and a labeled extension product larger than  
24 the standard is indicative of the presence of an insertion.

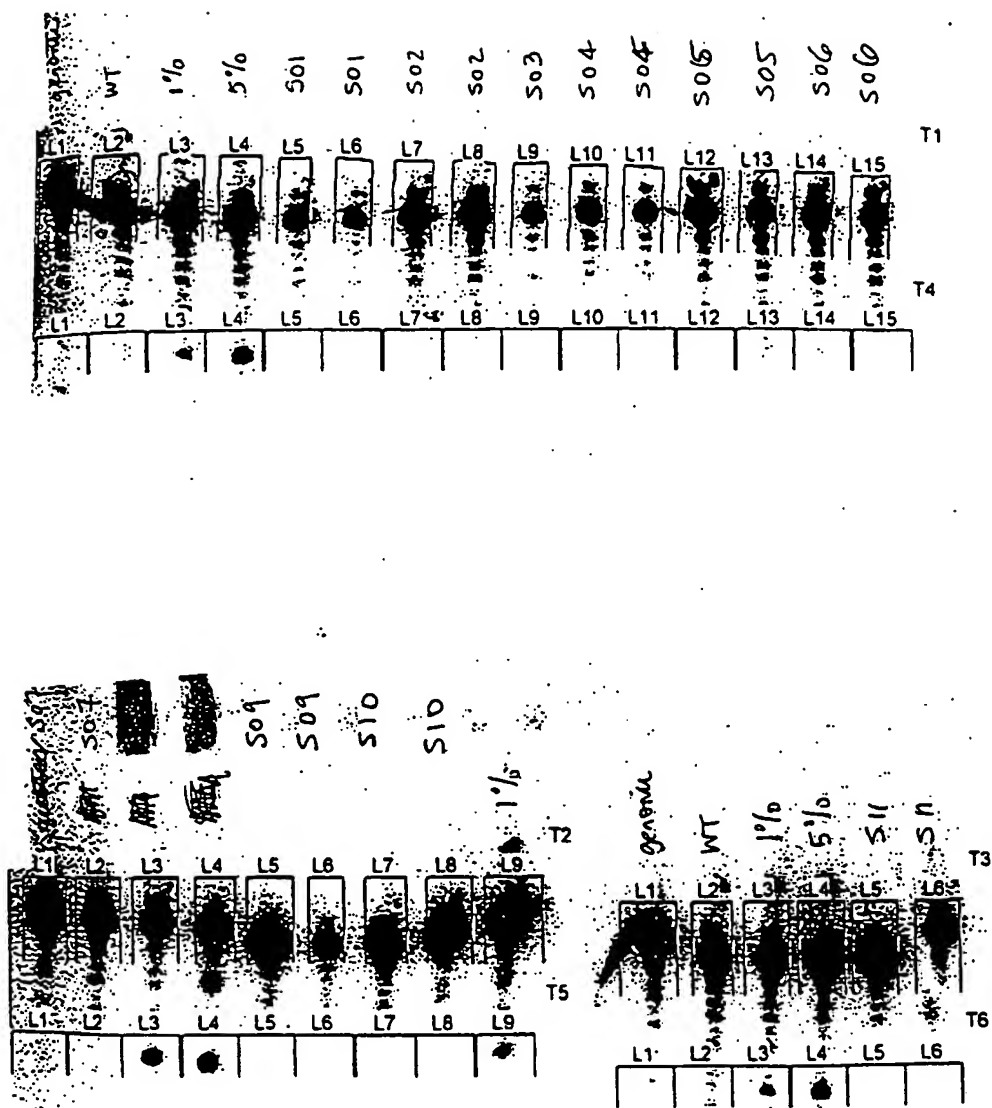
- 19 -

- 1 27. The method of claim 1 or 26, wherein the target region is greater than five bases  
2 long.
- 1 28. The method of claim 1 or 26, wherein the deletion or insertion is greater than three  
2 bases long.
- 1 29. The method of claim 1 or 26, wherein the standard is the wild-type labeled  
2 extension product.
- 1 30. The method of claim 1 or 26, wherein the standard is a molecular weight standard.

Figure 1A







**Figure 2**

## SEQUENCE LISTING

<110> Shuber, Anthony  
Pierceall, William  
Exact Laboratories, Inc.

<120> Methods for Detecting Mutations using Primer Extension

<130> EXT031PC

<140>

<141>

<150> US 09/371,991

<151> 1999-08-11

<160> 6

<170> PatentIn Ver. 2.0

<210> 1

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer 1

<400> 1

agcccttaac ctttttcagg

20

<210> 2

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Primer 2

<400> 2

gcccttaacc tttttcaggt

20

<210> 3

<211> 9

<212> DNA

<213> Homo sapiens

<220>

<223> wild-type APC sequence at codon 1309

<400> 3

gaaaagatt

9

<210> 4

<211> 5

<212> DNA

<213> Homo sapiens

<220>

<223> typical deletion at APC codon 1309



- 2 -

<400> 4  
gaaaa

5

<210> 5  
<211> 5  
<212> DNA  
<213> Homo sapiens

<220>  
<223> typical deletion found at APC codon 1309

<400> 5  
aaaag

5

<210> 6  
<211> 5  
<212> DNA  
<213> Homo sapiens

<220>  
<223> typical deletion found at APC codon 1309

<400> 6  
aaaga

5